# EXHIBIT 30

ground water

Issue Paper/

# Complexities in Hindcasting Models—When Should We Say Enough Is Enough?

by T. Prabhakar Clement

## Abstract

Groundwater models are routinely used in hindcasting applications to predict the past concentration levels in contaminated aquifers. These predictions are used in risk assessment and epidemiological studies, which are often completed either for resolving a court case or for developing a public-policy solution. Hindcast groundwater modeling studies utilize a variety of computer tools with complexity levels ranging from simple analytical models to detailed three-dimensional, multiphase, multispecies, reactive transport models. The aim of this study is to explore the value of using complex reactive transport models in hindcasting studies that have limited historic data. I review a chlorinated solvent exposure problem that occurred at a U.S. Marine Corp Base in Camp Lejeune, North Carolina and use it as an example to discuss the limits of hindcasting modeling exercises. The lessons learned from the study are used to reflect upon the following questions related to model complexity: How should we decide how much is enough? Who should decide when enough is enough?

## Introduction

On April 15, 2009, Professor Elizabeth Warren of Harvard Law School, formerly the chief of the congressional oversight panel for the troubled asset relief program, appeared on Jon Stewart's late night talk show to discuss our government's plan to stress-test failing banks. With a cynical smile on his face, Stewart asked: "How do you stress test a bank, if you will?" Professor Warren replied: "Well, you basically run it through a bunch of mathematical models and figure out whether the thing (bank) is financially healthy or the thing is really dead;" (note, in this context, she was just explaining the testing process). To my surprise, Stewart did not challenge this modeling effort and never asked a single follow-up question. The idea of conducting a computer-simulated stress test on a real bank, which is a complicated entity embedded within a dynamic economic web, should make anyone a cynic. Stewart's acceptance of this response without a question indicates the level of trust our society places in complex computer modeling efforts, especially when it is perceived that the efforts might provide benefits. This type of trust is not limited to economics. Scientists in other fields, including groundwater hydrology, tend to have such trust in models.

When critics challenge this trust in models, experts counter: What is the alternative? For the bank stress-test problem, we can, to some extent, answer this question by examining the opinions of some experienced investors in the financial industry. For example, on May 3, 2009, one of the world's renowned investors, Warren Buffet, dismissed the importance of the government stress tests in helping him assess banks (Bloomberg 2009). He stated: "I think I know their future, frankly, better than somebody that comes in to take a look." He also added that he judges banks by their "dynamism" and their "ability" to attract deposits.

Given these two vastly different positions (i.e., one based on computer modeling and another on expert opinion), one might wonder which one is more worthy for supporting a decision-making process. For the bank problem, one might have the following dilemma: Should I trust

Department of Civil Engineering, Auburn University, Auburn, AL 36849-5337; (334) 844-6268; fax: (334) 844-6290; clement@auburn.edu

Received June 2010, accepted September 2010.



doi: 10.1111/j.1745-6584.2010.00765.x

the complex mathematical assessment that purports to provide quantitative numbers for future deposits, profits, and earnings, or should I trust the qualitative conceptual assessment of an expert that provides subjective indicators such as a bank's "dynamism" or "ability"? As hydrogeologists, we are often challenged by our clients who want us to help them resolve complex policy questions or court cases. It is not easy to decide whether we should resolve such issues by using complex cutting-edge models or by using a balanced analysis of simpler calculations combined with expert judgments.

In this article, I present a philosophical discourse on the simplicity vs. complexity dilemma in the groundwater modeling field. There is an on-going debate on this topic (e.g., Haitjema 2006; Hill 2006; Gómez-Hernández 2006); however, past discussions have not considered the unique issues related to hindcasting efforts that use complex models. This article specifically focuses on a hindcast modeling case study that employed reactive transport models. The case study, which was recently reviewed by a National Academy panel administered by the National Research Council (NRC 2009), involves a chlorinated solvent contamination problem at a US Marine Corps Base in North Carolina. I served on this 14-member, interdisciplinary panel for more than 2 years and had an opportunity to review a wide range of health assessment, site characterization, and modeling studies. In this article, I will first provide a brief summary of the groundwater problem and will then use the lessons I have learned from this experience to reflect on the following two questions: How do we assess the required level of model complexity for a given hindcasting problem? Who should make the final decision about the complexity level?

## What Is a Hindcasting Model?

Mathematical models have been routinely used in the scientific literature to pursue epistemic research and/or policy research. The primary objective of an epistemic research effort is to create new knowledge that can help develop a mechanical scientific understanding of natural processes. The knowledge can then be used to generate testable hypotheses (predictions). A good example for an epistemic model is Einstein's general theory of relativity, which explained how gravity works and predicted, for example, the gravitational field would bend light. This prediction was later confirmed by Eddington's field data that documented the deflection of light by the sun's gravitational field from observations made during the solar eclipse in May 1919 (Dyson et al. 1920). The objective of policy-modeling efforts, on the other hand, is to provide "best possible estimates," which can be used by policy-makers to develop a timely decision to resolve a complex social problem that cannot be resolved using a mechanical scientific procedure. Policy models can be classified into forecasting models and hindcasting models. Forecasting models are used for predicting the future to resolve a potential problem. A good example of a forecasting exercise is the use of atmospheric models to predict the climate change effects.

Hindcasting models are used for predicting the past to understand and resolve historical problems. A good example of a hindcasting modeling application is the use of chemical fate and transport models to resolve public health issues related to a groundwater plume (e.g., the Woburn contamination problem; Bair and Metheny, 2011). Hindcasting applications are uniquely challenging because if we do not have the necessary past data for the system then there is no opportunity to collect the missing data. The scope of this article is limited to analyzing hindcasting policy models that employ complex reactive transport codes.

## Details of the Case Study

The case study considered here is based on a drinking water contamination problem that occurred in the 1950s and 1960s at a U.S. Marine Corp Base in Camp Lejeune (CLJ), North Carolina. The base is a 246-square-mile military training facility located in Onslow County, southeast of the City of Jacksonville, North Carolina. The site has multiple contaminated areas that are impacted by several types of hazardous chemicals. In this article, I will focus on a tetrachloroethylene (PCE) plume present in the Tarawa Terrace (TT) area (Figure 1) for which extensive modeling information is available (Maslia et al. 2007; NRC 2009). The PCE plume originated from an off-site dry cleaning facility, ABC One Hour Cleaners, which started operation in 1953 (Figure 1). The site has a considerable amount of hydrogeological characterization data, but limited chemical/biological characterization data. Details of the site characterization data available are discussed in Harden et al. (2004), Maslia et al. (2007), and Faye and Green (2007).

The groundwater contamination problem was first discovered in the early 1980s when a routine water quality survey indicated the presence of unknown organic compounds in the drinking water. Further investigations revealed that the water supplied by the on-site water treatment plant, which extracted water from multiple wells installed in the local aquifer (see Figure 1 for well locations), was contaminated with PCE and its degradation products. Later, it was determined that the drinking water was also contaminated with other chemicals including petroleum products. As the modeling studies completed so far have solely focused on the PCE contamination problem, I will limit the discussions to chlorinated solvent-related issues.

At the CLJ site, it is estimated that more than a million people have been exposed to the contaminated water delivered between the mid-1950s and the mid-1980s. Currently, more than 156,000 people have formally registered with the Marines Corps to get more information about the contamination. Several former CLJ residents have moved forward with claims against the Marine Corps complaining that the contaminated water has caused a variety of cancers and other ailments. To address these complaints,

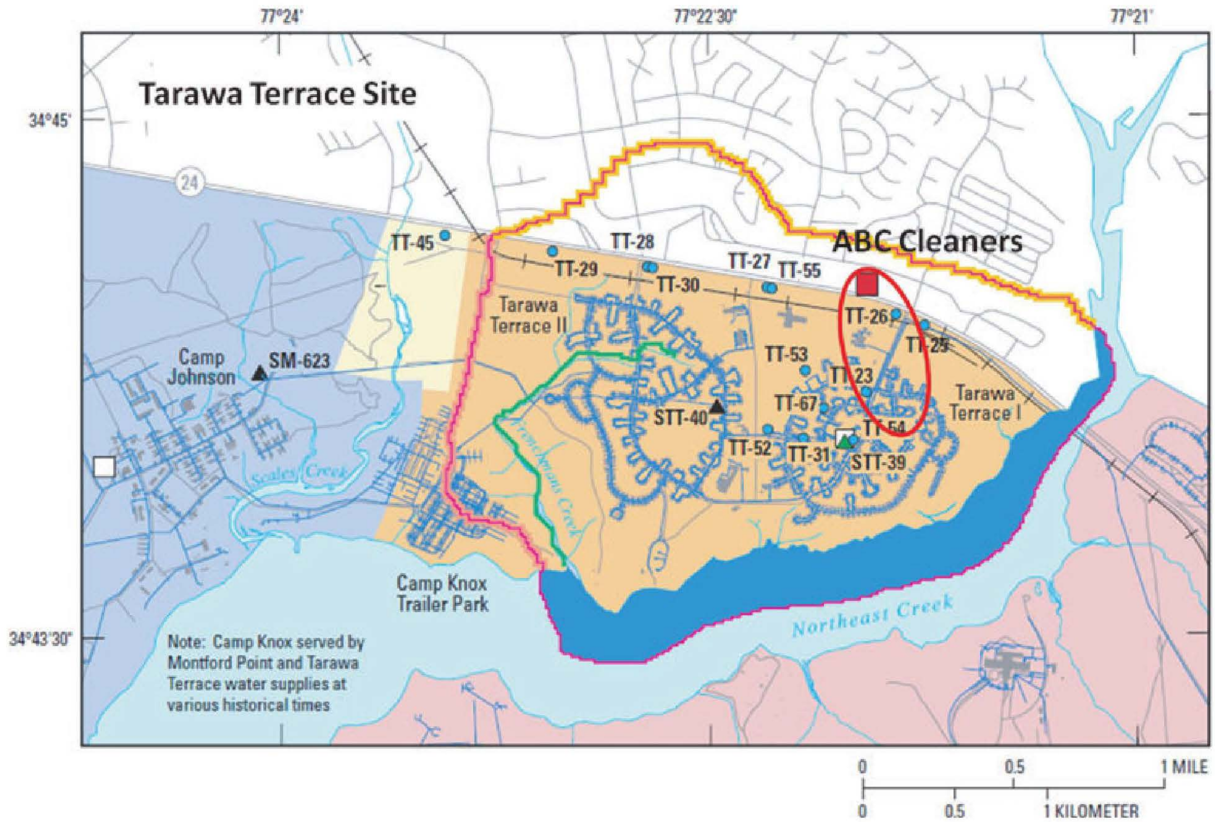


**Figure 1. Details of the Tarawa Terrace site. Pumping wells are indicated by blue dots labeled with letters TT. The red square indicates the location of ABC Cleaners. Triangles are storage tanks and are not relevant to this discussion (from Maslia et al. 2007).**

the Agency of Toxic Substances and Disease Registry (ATSDR) conducted a study that examined the association between well-defined, quantitative levels of PCE and TCE (trichloroethylene) in drinking water and the risk of developing specific birth defects—spina bifida, anencephaly, cleft lip, cleft palate, childhood leukemia, and non-Hodgkin's lymphoma (Maslia et al. 2007). The study included groundwater modeling efforts to reconstruct the past contamination scenarios and also interviews to obtain residential history, information on water consumption habits, and other risk factors. ATSDR postulated that by using model-derived drinking water concentrations and the interview data, associations between exposure to PCE and TCE and the risk of particular health outcomes could be thoroughly examined (Maslia et al. 2007). ATSDR used the public-domain codes MODFLOW and MT3DMS to predict the fate and transport of PCE, and an advanced research code TechFlowMP (Jang and Aral 2008), to predict the concentrations of PCE along with its degradation byproducts TCE, *trans*-1,2-dichloroethylene (*trans*-DCE), and vinyl chloride (VC).

Based on the modeling studies, researchers reconstructed the historical contamination levels and the model-predicted concentrations were widely disseminated to various groups. Figures 2 and 3 show example modeling results that were made available to scientists interested in conducting exposure assessment studies. Figure 2 shows probabilistic MT3DMS predictions for the historic PCE concentration levels in the drinking water, generated using the Monte Carlo approach. The range of PCE concentrations derived from the probabilistic analysis (shown as a band in the figure) represents 95% of all possible results. These values were derived from multiple realizations of the MT3DMS model runs. Figure 3 shows the results from a multispecies, multiphase research code TechFlowMP, which was used to predict the historic concentrations of the biodegradation byproducts TCE, *trans*-DCE (*cis*-DCE was not considered), and VC. It is important to note that, as shown in Figure 2, the model was calibrated to limited number of data points, which are PCE levels measured in finished water samples collected in the early 1980s. Also, note that Figure 3 does not report any measured data for the biodegradation products TCE, DCE, or VC.

These model results were presented to former CLJ residents (via websites, public meetings, and reports), health scientists, and congressional committees. All three groups appear to have accepted the results and the modeling methodologies. The results appeared to be reasonable because the Monte Carlo simulations indicated a narrow band within which 95% of the model-simulation results resided. The figure shows that the 95% confidence band becomes narrower as we move back from the 1980s (where there is no data); this implies that the groundwater model was able to make confident hindcasts

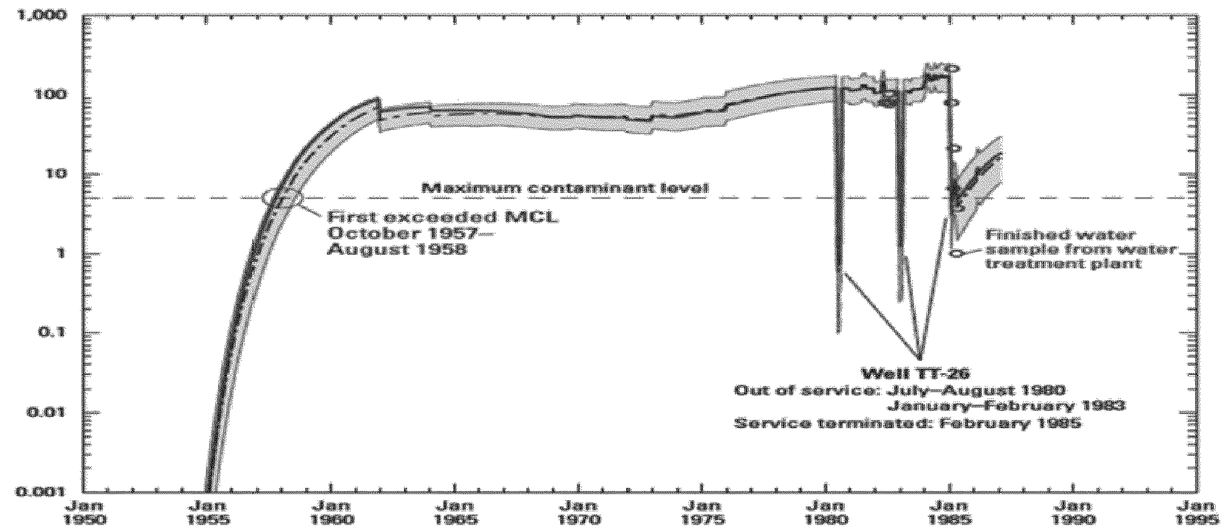


**Figure 2. Predicted concentration levels of PCE (μg/L) in the finished water delivered by the Tarawa Terrace treatment plant. MT3DMS model results. The center line is the mean concentration, upper limit is the 97.5% and lower limit is the 2.5% of 510 Monte Carlo simulations (from Maslia et al. 2007).**

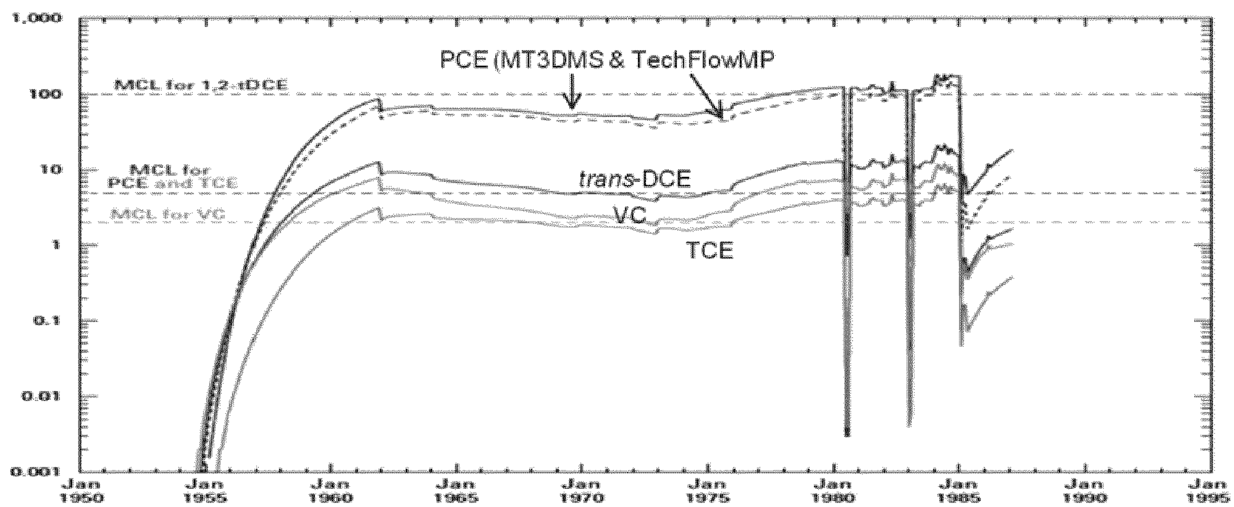


**Figure 3. Predicted concentration levels of PCE, TCE, *trans*-DCE, and VC (μg/L) in the finished water delivered by the Tarawa Terrace water treatment plant. TechFlowMP model results (from Maslia et al. 2007).**

for the 1950s and 1960s even if there are no past data to calibrate the model. The figure also shows that closer to the initial starting point the confidence band is almost 100%, implying that our knowledge of initial conditions, initial source loadings, and initial stresses is almost exact. Figure 3 shows the model-predicted values of various PCE biodegradation products. Health scientists found these TechFlowMP predictions to be useful because they provided quantitative concentration estimates for assessing the health impacts of more toxic biodegradation products such as VC. Based on some favorable feedbacks, researchers planned follow-up efforts to conduct additional modeling studies to make hind-casting predictions for other contaminated areas including Hadnot Point and Holcomb Boulevard, which are located within the CLJ site. One of the tasks of the NRC (2009) effort was to review these proposals and make impartial recommendations for future groundwater studies.

## Constraints on Complexity Due to Process Uncertainties and Data Limitations

The overall objective of the CLJ study was to determine whether exposure to the contaminated water caused the reported health problems. The exposure assessment efforts not only required contaminant concentration levels, but also other site-specific historic data such as total water usage by various impacted residents and their daily water consumption patterns, which are possibly unknowable information. Furthermore, even if one had a "perfect" groundwater model, the final outcomes of the study would

have considerable uncertainties due to lack of knowledge about actual exposures, their impacts on human health, and the difficulty of making causal inference from observational studies. Oreskes (1998) identified four possible limitations related to exposure/health assessment studies, which arise from theoretical, empirical, parametrical, and temporal uncertainties. Theoretical uncertainties are related to processes which we simply do not understand and hence do not have the correct theoretical (mathematical) description to model the process. Empirical uncertainties are related to factors that we cannot measure. This would include having limited resources to collect samples (e.g., blood samples of the exposed population) and analytical uncertainties in quantifying low levels of chemicals in tissues or blood samples. Parametric uncertainties are the errors introduced when we reduce complex phenomena to a single (fixed or varying) input parameter. Temporal uncertainties arise from the assumption that systems are stable in time. Oreskes argued that from a biological standpoint systems may not be stable in time; for example, high and low levels of blood concentration could be as important as the mean and might affect human health in ways that are neither fully understood nor fully measured.

The groundwater modeling field also has several issues that are quite similar to those pointed out by Oreskes. Over the past three decades, the groundwater modeling community has progressed considerably in addressing these issues. The basic theoretical framework for simulating flow and nonreactive transport is now reasonably well understood and has been routinely used for analyzing practical problems. Powerful analytical and numerical approaches are now available for efficiently solving groundwater problems. The analytical advances have led to the development of efficient close-form solutions to various reactive transport problems and public-domain screening tools (Aziz et al. 2000; Clement et al. 2002; Quezada et al. 2004; Srinivasan and Clement 2008). The numerical advances have led to the development of well-accepted public-domain codes such as MODFLOW and MT3DMS, and related reactive transport codes such as RT3D (Clement et al. 1998) and PHT3D (Prommer et al. 2003). In addition, calibration/uncertainly assessment tools such as PEST (Doherty 2005) and UCODE (Poeter et al. 2005) have also received widespread acceptance. Despite these advances, contaminant transport models, especially reactive transport models used for simulating the fate and transport of solvent plumes evolving from dense nonaqueous phase liquid (DNAPL) sources, still have several important limitations. Over 20 years ago, Anderson (1983) reviewed the state-of-the-art of groundwater modeling practices and warned: "Be careful! The Emperor has no clothes!" Hunt and Welter (2010) pointed out that complex groundwater models will always have structural (or theoretical) errors, also known as "unknown unknowns." More recently, Konikow (2011) reviewed the state of solute transport modeling and concluded that the secret to a successful solute transport modeling effort is simply to lower expectations.

Bioreactive transport problems involving DNAPL contaminants, such as PCE, often require model formulations that involve multiple parameters which make predictions more problematic. While recent research efforts have advanced our understanding of biological processes related to chlorinated solvents (McCarty 1997; Bradley et al. 2008), quantitative prediction of PCE biodegradation using reactive transport models is still beyond the state of standard practice. Given this state of knowledge, it is worth examining the value of CLJ hindcasts, which were derived from complex bioreactive transport models that were fitted to few data points, for developing a policy solution to the problem.

One of the important concerns that limit the use of bioreactive transport models at chlorinated solvent sites is the lack of problem-specific information on input parameters. A key input to any transport model is information related to the source. Unfortunately, this information is one of the most unreliable types of input deduced from qualitative assessments. This is especially true in hindcasting applications involving DNAPL wastes. At the TT site, the contaminant of concern, PCE (a DNAPL), was sporadically disposed by a dry cleaner in the DNAPL form, along with other waste products, into a septic tank. Site characterization data indicated that a shallow monitoring well installed close to the dry cleaning facility recorded an extremely high PCE concentration of 12,000 µg/L (Faye and Green 2007). Such high-concentration levels would indicate that the source region might still have residual DNAPL. At DNAPL-contaminated source regions spatial variability in mass is almost inevitable and consequently the mass detection process will be extremely difficult and uncertain (Abriola 2005). Detailed modeling of PCE migration processes from the septic tank requires input data related to waste disposal practices, historical infiltration levels, unsaturated zone properties, effective solubility level of the mixed-waste DNAPL, and its dissolution kinetics. In summary, the way (what, when, where, and how) PCE was discharged into the system and how long the PCE waste resided in DNAPL form are important factors controlling historic plume concentrations. Yet this critical past information cannot be obtained.

The TT water supply system extracted groundwater from multiple wells installed in a highly heterogeneous, multilayer aquifer. These wells were operated in a cyclic manner. The influent concentrations of degradation species (such as TCE, DCE, or VC) would have depended on the location of the pumped well from which the water was extracted at a given time and the level of subsurface microbial activity at that location during that time period. Literature data show that subsurface microbial reactions can be mediated by a complex set of biogeochemical mechanisms that are facilitated by a variety of microbes (McCarty 1997; Clement et al. 2000; Bradley et al. 2008). Microbiologists are still debating whether a specialized microbial species, such as *Dehalococcoides*, or a variety of natural microbial populations would facilitate degradation of chlorinated solvents (Major et al. 2003; Nyer et al. 2003). They are also debating, to some extent,

whether the degradation byproduct DCE will be in *cis*-form or in *trans*-form (Miller et al. 2005), although most field studies have shown that DCE is predominately present in the *cis*-form (Wiedemeier et al. 1999).

It is now well established that reductive dechlorination reactions are limited by the availability of a degradable carbon source that can supply hydrogen (McCarty 1997; Yu et al. 2005; Bradley et al. 2008). However, it is difficult to accurately simulate this limitation in large-scale field problems that have multiple competing biogeochemical processes. Clement et al. (2000) proposed a reaction-zone approach to incorporate carbon limitations indirectly at a chlorinated solvent field site in Delaware, USA. Rolle et al. (2008) proposed a kinetic framework for modeling the interactions between carbon and terminal electron acceptors at a landfill site in Italy. However, these are research models that require extensive field-measured biogeochemical data.

Accurate reconstruction of biodegradation byproducts in the drinking water requires historical data for groundwater pumping rates, pumping patterns (recall that the TT treatment plant extracted groundwater from multiple wells in a cyclic manner), geochemical data, concentrations of microbial populations, microbial growth kinetics, and secondary removal rates within treatment units and pipelines. This would necessitate compilation of an enormous amount of past information, most of which is very likely not available at the TT site.

## How Should We Decide How Much Is Enough?

The above discussions illustrate the inherent difficulties in developing a bioreactive transport model for reconstructing the PCE contamination scenarios at the TT site that occurred 30 to 40 years ago. Given these difficulties, for hindcasting applications such as the CLJ study, it is perhaps prudent to limit the required level of model complexity to a level that is consistent with the level of available data. This recommendation is not new; it is simply Occam's razor, a well-accepted principle that advocates model parsimony (Hill 2006, NRC 2007). This is a logical approach that necessitates the use of simple models when we have limited data. This practice is particularly more appropriate for hindcasting modeling exercises where it is virtually impossible to obtain missing historical data.

In the literature, researchers have criticized such simplistic modeling approaches, though most of the criticisms were developed in the context of model use in epistemic or forecasting applications. For example, Cunge (2003) argues that simple models add the certainty of a poor quality of modeling to the data uncertainty, and the synergy of the two is likely to result in a very poor representation of reality. He recommends that in a true good practice, a lack of adequate data necessitates the use of the most advanced and reliable modeling tools. It is important to note that Cunge's discussions were aimed toward epistemic (not policy) modeling exercises. Also, his recommendations assume that available "advanced" models would be able to realistically simulate all natural processes.

Oreskes (2003) noted that we tend to have more intuitive faith in complex models because they allow us to simulate more processes. However, as we add more processes (and parameters) to a model, the overall certainty of its predictions might decrease. Ironically, the "truer" the model, the more difficult it is to show that it is "true" (Oreskes 2003). Modeling critics have also pointed out several case studies to illustrate the failure of complex models at various levels; they have argued that mathematical complexities have little value in predicting the behavior of natural systems (Pilkey and Jarvis 2007). Complex computer models are based on reductionism, which assumes one can decompose natural complexity into simple components at an appropriate scale. Rigler and Peters (1995) critiqued such approaches and stated that computers gave reductionists the tools to approach an ecosystem as the sum of its parts, which leads to the conclusion that these tools are inadequate. Chave and Levin (2003) concluded that natural ecological processes (e.g., activity of microbial systems) are not only complex but are also adaptive; moreover, there is no single correct scale on which to study their dynamics.

Recently, the problem of equifinality in complex systems has been discussed extensively in the hydrological literature (Beven 1993). Equifinality is the recognition that different initial states, model structures, and/or parameter sets can lead to similar end states. For the CLJ problem, for example, the site only had a limited number of PCE data points, which were short-term averaged random grab measurements made in the early 1980s (Figure 2). The calibration exercises were aimed toward fitting the monthly-averaged model predictions to these limited data points, within a predefined fixed target level, with an assumption that the calibrated model would be able to hindcast the historical levels of PCE and its byproducts in the 1950s, 1960s, and 1970s. However, due to limitations in our understanding of natural processes and due to inaccuracies in measurement methods, several complex models with many different model structures and initial conditions might fit these limited observations equally well. Beck (1987) reviewed various water quality modeling methods and concluded that a lack of model identifiability has been an outstanding difficulty in the interpretation of observed system behavior and there is ample evidence to show that larger (more complex) models are easily capable of generating highly imprecise predictions. For the TT site, due to sparsity of observations, it will be difficult to identify a unique (or precise) model structure.

Given these limitations, numerical modeling approaches used in data limited, hindcasting policy applications should perhaps employ simpler conceptual approaches that use lower dimensionality (e.g., depth-averaged models), average flow, simpler reactions, spatially-averaged model parameters, temporally-averaged source loading patterns, to name a few. In addition, one could also use simpler tools such as analytical models

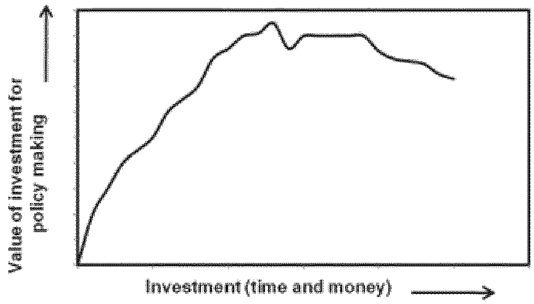


**Figure 4. Conceptual relationship between modeling investment and its value for developing policy solutions.**

(Haitjema 2006), conceptual calculations that are based on available data (Bredehoeft 2003), and approximations derived using mass-balance calculations. As summarized in NRC (2009), the level of uncertainty associated with such simpler models perhaps will not be lower than the complex predictions; however, these simpler models could be completed quickly in a cost-effective manner and this should help speed up the decision-making process. Simpler models could also provide opportunities to communicate the conceptual limitations of the hindcasting exercise more effectively to the broader user group (e.g., health scientists, politicians, and the concerned public). Denton and Sklash (2006) reviewed several case studies of model applications in court rooms and pointed out that complicated models add little value and could create more opportunities for confusion and challenges.

Figure 4 illustrates a hypothetical conceptual relationship between model investment and its value for aiding the public-policy development process. As shown in the figure, the initial investments made in a modeling effort can help develop a better understanding of the contamination scenario that can be useful for the decision-making process. However, the return on the investment might quickly become marginal and the value of the information gained from new studies would approach a plateau in a nonlinear manner. More importantly, it is possible that larger investments (to develop complex models or to conduct advanced scientific studies) could considerably delay and complicate the decision-making process and even have a negative impact. This is particularly true for hindcasting applications where higher levels of complexity could simply muddle the issues and make the decision-making process a lot more political, thus hampering the process rather than aiding it.

Model development is a dynamic exercise and it is difficult to complete an a priori assessment of model worthiness (i.e., its value related to decision making) at various investment levels for a given project. Therefore, the level of investment (in model complexity) needed for a problem is necessarily a subjective judgment that should be made after carefully considering the information related to available site data, available resources (time and money), and the modeling objectives. The conceptual relationship illustrated in Figure 4 can be a useful guide for integrating all the information to develop a balanced level of model complexity for a given problem. It can serve as a mental model that can help answer the rhetorical question—How much is enough?

## Who Should Decide When Enough Is Enough?

Most practical modeling studies are performed by consulting companies that rarely use cutting-edge research tools. Years of model use in litigation efforts have made it clear that using research codes on high-visibility projects is not a good idea. However, in some cases, scientific teams with certain gravitas might convince agencies to support the application of their cutting-edge tools. In such cases, advocating the use of appropriate simpler tools is not easy for the experts who are performing the work. Jamieson (2000) pointed out that scientists live in a highly competitive environment where funding for research is limited. Involvement in policy-modeling projects helps scientists present themselves as real-world problem solvers, which helps secure funding for their scientific pursuits. Sarewitz and Pielke (2000) stated that advocating the use of advanced cutting-edge models is always an attractive short-term solution because it benefits not only the scientists who receive the funding, but also the politicians who fund their effort. It is a "win-win" strategy where the scientists receive direct funding to develop and test their latest tools, and politicians can point to these "scientific" projects as actions and safely defer making difficult decisions as they wait for the study results. Moreover, concerned citizen groups feel good about such scientific pursuits as they believe that the scientists and politicians are doing their best to resolve their problem. In the end, all three parties tend to rationalize the decision and convince themselves that they are doing the right thing. Hence, it will be difficult, if not impossible, for these interest groups to make an impartial judgment call on the required level of model complexity. Use of external peer reviewers, who have little or no self-interest in the project, would perhaps be the more appropriate option.

For the CLJ project, the judgment call was made by the NRC panel, which consisted of a diverse group of 14 experts who volunteered their time to study various aspects of the problem for 2 years and prepared a report, which was reviewed by 10 external reviewers. The panel made the following conclusions (NRC 2009):

> the Tarawa Terrace and Hadnot Point supply systems were contaminated with volatile organics, particularly PCE, TCE and DCE, for decades ending in the middle of 1980s (p. 64). There were divergent views among the committee members about the probability that each would assign to whether adverse health effects have in fact occurred, but there was consensus among them that scientific research is unable to provide more definitive answers (p. 22). [This implied that] it cannot be determined reliably whether diseases and disorders experienced by former residents and workers at Camp Lejeune are associated with their exposure to contaminants in the water supply because of

> data shortcomings and methodological limitations, and these limitations cannot be overcome with additional study. Thus, the committee concludes that there is no scientific justification for the Navy and Marine Corps to wait for the results of additional health studies before making decisions about how to follow up on the evident solvent exposures on the base and their possible health consequences. The services should undertake the assessments they deem appropriate to determine how to respond in light of the available information (p. 13).

The panel also recommended:

> the use of simpler approaches (such as analytic models, average estimates based on monitoring data, mass-balance calculations, and conceptually simpler MODFLOW/MT3DMS models) that use available data to rapidly reconstruct and characterize the historical contamination (p. 65). Also, policy changes or administrative actions that would help to resolve the controversy should proceed in parallel with the studies (if they are continued) rather than in sequence (p. 22).

As voluntary expert committees, such as the NRC panel, do not have any direct self-interest, their collective wisdom is likely to recommend a reasonable practical solution, although by no means would it be the perfect solution.

The overall response to the NRC study was mixed. Various groups of health scientists, environmental activists, one of the modeling teams, and the former CLJ residents were disappointed and severely criticized the study's conclusion that additional scientific studies cannot provide more definitive answers. In 2009, two senators from North Carolina introduced a bill to furnish hospital care, medical services, and nursing home care to veterans who were stationed at the base while the water was contaminated. In February 2010, a North Carolina congressman introduced *The Janey Ensminger Act* in the House of Representatives to require the Department of Veterans Affairs to provide the healthcare benefits. These new policy developments directly address the healthcare needs of the community. The lead government agency, ATSDR, developed a professional response to the NRC study that included the following statements (ATSDR 2009):

> ATSDR will apply simpler modeling techniques for Hadnot Point and Holcomb Boulevard than those used for Tarawa Terrace. The Hadnot Point area is significantly larger than the Tarawa Terrace area and contains multiple contaminant source locations. Applying the complex numerical models used at Tarawa Terrace to the entire Hadnot Point and Holcomb Boulevard areas would be time consuming, costly, and add another level of uncertainty to the water-modeling analyses.

ATSDR's (2009) plan included continuation of some of the epidemiology studies, as they viewed these studies will be scientifically useful (will have epistemological value), and will also be helpful to the community of service men and women and their families (an important social value).

## Concluding Remarks

Reactive transport models are useful tools that can help us gain insights into the importance of key biological and/or chemical variables and their causes and effects. As most of us are limited by our linear thinking, there is always a conceptual gap in our understanding of how all the various parts of a nonlinear biochemical system couple with transport and respond as the whole. Reactive transport models can help fill this gap by integrating an enormous amount of diverse information (physical, chemical, and biological data) into a unified rational framework. The simulation results can be used to construct reasonable qualitative arguments as to why certain processes or events can or cannot occur. However, it is important that we understand the limits of these tools and recognize that they are better viewed as computer-aided thinking tools rather than computer-aided prediction tools.

While critiquing the mathematical models used in sociology and economics, George Andrews, the current President of the American Mathematical Society, made the following statement (Andrews 1988): "Mathematics is not a mysterious substitute to educated common sense. When mathematics is abused—used in areas where measurements are extremely difficult or impossible—it is, at best, a nuisance and, at worst, a trick to disguise ulterior motives." If we simply replaced the word "mathematics" with "groundwater model," then these wise words would literally transform into a prophetic statement about groundwater modeling! Victor Baker, the former President of the Geological Society of America, said "allowing the public to believe that a problem can be resolved . . . through elegantly formulated . . . models is the moral equivalent of a lie" (Pilkey and Jarvis 2007, p. 188).

When debating the worthiness of hindcast modeling efforts that have direct implications on public policy, it is difficult to say whether it is a scientific debate or a moral debate. In such instances, it is perhaps worth reflecting upon some of the wise statements made by our scholarly peers. Such reflections might inspire us to raise a simple self-assessment question: Is this a worthy effort for developing a sound public-policy decision? It is extremely difficult to give an honest answer to this question, especially when personal interests are at stake. As scientists, we all suffer from some level of cognitive dissonance and have an uncanny ability to rationalize our futile efforts. However, we, the scientific community, owe an honest answer to our fellow citizens who put enormous faith in our abilities and fund us to explore the beauty of natural systems and their mystifying connection to mathematics. Also, I believe that such a self-assessment should provide the ultimate wisdom for understanding the worthiness of our scientific pursuits and would guide us when to say enough is enough. But I admit that accepting and acting upon this self-assessment is a lot easier said than done!

## Acknowledgments

I would like to thank Mary Anderson for encouraging me to write this article and for providing editorial reviews.

My sincere thanks to Frank Schwartz for his detailed reviews; two other anonymous reviewers for the journal *Ground Water* also provided useful comments that helped strengthen this article. This is an unfunded effort; I am grateful to the Department of Civil Engineering, Auburn University, for supporting part of my past 3 years of research time to reflect on this esoteric topic. I benefitted from the discussions with NRC staff members and the committee members who served with me on the Camp Lejeune NRC panel. Special thanks to the Committee Chair, David Savitz and to the exposure Subcommittee Chair, John Nuckols; their challenging questions helped organize my thought process. Several of my colleagues including M.J. Simpson, R.R. Goswami, J. Vasconcelos, A. Son, M.O. Barnett, F.J. Molz, H. Abbas, L.E. Sheppard, J.S. Hayworth, and G. Jeppu reviewed this manuscript and provided valuable suggestions. Any errors of fact or opinion are, however, solely mine and not of any reviewer, agency, committee, or panel.

## References

Abriola, L.M. 2005. Contaminant source zones: Remediation or perpetual stewardship? *Environmental Health Perspectives* 113, no. 7: A438–A439.

Anderson, M.P. 1983. Ground-water modeling—The emperor has no clothes. *Ground Water* 21, no. 6: 666–669.

Andrews, G.E. 1988. Outlook: Numbers don't lie, but . . . , Research Penn State, Penn State University, 3.

ATSDR. 2009. Contaminated drinking water and health effects at Marine Base Camp Lejeune: Final plans of the Agency for Toxic Substances and Disease Registry. http://www.atsdr.cdc.gov/sites/lejeune/ (accessed June 5, 2010).

Aziz, C.E., C.J. Newell, J.R. Gonzales, P. Haas, T.P. Clement, and Y. Sun. 2000. BIOCHLOR v1.0—Natural attenuation decision support system. User's Manual, United States Environmental Protection Agency Report, Cincinnati, Ohio, EPA/600/R-00/008.

Bair, E.S., and M.A. Metheny. 2011. Lessons learned from the landmark "*A Civil Action*." *Ground Water* (this issue).

Beck, M.B. 1987. Water quality modeling: A review of the analysis of uncertainty. *Water Resources Research* 23, no. 8: 1393–1442.

Beven, K.J. 1993. Prophecy, reality and uncertainty in distributed hydrological modeling. *Advances in Water Resources* 16, 41–51.

Bloomberg. 2009. Buffett dismisses government stress tests. http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aOBs7wYngdiY (accessed September 16, 2010).

Bradley, P.M., F.H. Chapelle, and F.E. Löffler. 2008. Anoxic mineralization: Environmental reality or laboratory artifact? *Ground Water Monitoring and Remediation* 28, no. 1: 47–49.

Bredehoeft, J. 2003. From models to performance assessment: The conceptualization problem. *Ground Water* 41, no. 5: 571–577.

Chave, J., and S. Levin. 2003. Scale and scaling in ecological and economic systems. *Environmental and Resource Economics* 26, 527–557.

Clement, T.P., Y. Sun, B.S. Hooker, and J.N. Petersen. 1998. Modeling multi-species reactive transport in groundwater aquifers. *Groundwater Monitoring & Remediation Journal* 18, no. 2: 79–92.

Clement, T.P., C.D. Johnson, Y. Sun, G.M. Klecka, and C. Bartlett. 2000. Natural attenuation of chlorinated solvent compounds: Model development and field-scale application. *Journal of Contaminant Hydrology* 42, 113–140.

Clement, T.P., M.J. Truex, and P. Lee. 2002. A case study for demonstrating the application of U.S. EPA's monitored natural attenuation screening protocol at a hazardous waste site. *Journal of Contaminant Hydrology* 59, no. 1–2: 133–162.

Cunge, J.A. 2003. Of data and models. *Journal of Hydroinformatics* 5, 75–98.

Denton, C.M., and M.G. Sklash. 2006. Contaminant fate and transport in courtroom. In *Contaminated Soils, Sediments, and Water—Success and Challenges*, ed. E.J. Calabrese, P.T. Kostecki, and J. Dragun, 81–118. New York: Springer. DOI: 10.1007/0-387-28324-2.

Doherty, J. 2005. *PEST Model-Independent Parameter Estimation, User Manual*, 5th ed. Watermark Numerical Computing.

Dyson, F.W., A.S. Eddington, and C.R. Davidson. 1920. A determination of the deflection of light by the sun's gravitational field, from observations made at the solar eclipse of May 29, 1919. *Philosophical Transactions of Royal Society, Series A* 220, 291–333. doi: 10.1098/rsta.1920.0009.

Faye, R.E., and J.W. Green. 2007. *Chapter B: Geohydrologic Framework of the Castle Hayne Aquifer System.* Atlanta, GA: Agency for Toxic Substances and Disease Registry.

Gómez-Hernández, J.J. 2006. Complexity. *Ground Water* 44, no. 6: 782–785.

Haitjema, H. 2006. The role of hand calculations in ground water modeling. *Ground Water* 44, no. 6: 786–791.

Harden, S.L., S.S. Howe, and S. Terziotti. 2004. Direction of ground-water flow in the surficial aquifer in the vicinity of impact areas G-10 and K-2, Marine Corps Base Camp Lejeune, North Carolina. U.S. Geological Survey Scientific Investigations Report 2004-5270. U.S. Department of the Interior, U.S. Geological Survey [online]. http://pubs.usgs.gov/sir/2004/5270/pdf/report.pdf (accessed August 15, 2010).

Hill, M.C. 2006. The practical use of simplicity in developing ground water models. *Ground Water* 44, no. 6: 775–781.

Hunt, R.J., and D.E. Welter. 2010. Taking account of "unknown unknowns." *Ground Water* 48, no. 4: 477.

Jamieson, D. 2000. Prediction in society. In *Predictions*, ed. D. Sarewitz, R.A. Pielke, and R. Byerly, 315–325. Washington, DC: Island Press.

Jang, W., and M.M. Aral. 2008. *Chapter G: Simulation of Three-Dimensional, Multi-species, Multi-phase Mass Transport of Tetrachloroethylene (PCE) and Associated Degradation By-products.* Atlanta, Georgia: Agency for Toxic Substances and Disease Registry.

Konikow, L.F. 2011. The secret to successful solute-transport modeling. *Ground Water*. In press.

Major, D., E. Edwards, P.L. McCarty, J. Gossett, E. Hendrickson, F. Loeffler, S. Zinder, D. Ellis, J. Vidumsky, M. Harkness, G. Klecka, and E. Cox. 2003. Discussion of environment vs bacteria or let's play "Name that Bacteria." *Ground Water Monitoring & Remediation Journal* 23, no. 2: 32–48.

Maslia, M.L., J.B. Sautner, R.E. Faye, R.J. Suárez-Soto, M.M. Aral, W.M. Grayman, W. Jang, J. Wang, F.J. Bove, P.Z. Ruckart, C. Valenzuela, J.W. Green Jr., and A.L. Krueger. 2007. *Analyses of Groundwater Flow, Contaminant Fate and Transport, and Distribution of Drinking Water at Tarawa Terrace and Vicinity—Chapter A: Summary of Findings.* Atlanta, Georgia: Agency for Toxic Substances and Disease Registry.

McCarty, P.L. 1997. Breathing with chlorinated solvents. *Science* 276: 1521.

Miller, G.S., C.E. Milliken, K.R. Sowers, and H.D. May. 2005. Reductive dechlorination of tetrachloroethene to trans-dichloroethene and cis-dichloroethene by PCB-dechlorinating bacterium DF-1. *Environmental Science and Technology* 39, no. 8: 2631–2635.

NRC. 2007. Models in Environmental Regulatory Decision Making. Board on Environmental Studies and Toxicology, National Academies Press, 286.

NRC. 2009. Contaminated Water Supplies at Camp Lejeune: Assessing Potential Health Effects. Board on Environmental Studies and Toxicology, National Academies Press, 328.

Nyer, E.K., F. Payne, and S. Suthersan. 2003. Environment vs. bacteria or let's play "Name that Bacteria." *Ground Water Monitoring & Remediation* 23, no. 1: 36–45.

Oreskes, N. 1998. Evaluation (not validation) of quantitative models. *Environmental Health Perspective* 106, no. 6: 1453–1460.

Oreskes, N. 2003. The role of quantitative models in science. In *Models in Ecosystem Science*, ed. C.D. Canham, J.J. Cole, and W.K. Lauenroth, Chapter 2, 13–31. New Jersey: Princeton University Press.

Pilkey, O.H., and L.P. Jarvis. 2007. *Useless Arithmetic—Why Environmental Scientists Can't Predict the Future*, 230. New York: Columbia University Press.

Poeter, E.P., M.C. Hill, E.R. Banta, S. Mehl, and S. Christensen. 2005. UCODE_2005 and Six Other Computer Codes for Universal Sensitivity Analysis, Calibration, and Uncertainty Evaluation: U.S. Geological Survey Techniques and Methods 6-A11, 283.

Prommer, H., D.A. Barry, and C. Zheng. 2003. MODFLOW/MT3DMS based reactive multi-component transport modeling. *Ground Water* 41, no. 2: 247–257.

Quezada, C.R., T.P. Clement, and K.K. Lee. 2004. Generalized solution to multi-dimensional, multi-species transport equations coupled with a first-order reaction network involving distinct retardation factors. *Advances in Water Resources* 27, 507–520.

Rigler, F.H., and R.H. Peters. 1995. *Science and Limnology.* Germany: Ecology Institute, Oldendorf/Luhe.

Rolle, M., T.P. Clement, R. Sethi, and A.D. Molfetta. 2008. A kinetic approach for simulating redox-controlled fringe and core biodegradation processes in groundwater: Model development and application to a landfill site in Piedmont, Italy. *Hydrological Processes* 22, no. 25: 4905–4921.

Sarewitz, D., and R.A. Pielke. 2000. Prediction in science and policy. In *Predictions*, ed. D. Sarewitz, R.A. Pielke, and R. Byerly, 11–22. Washington, DC: Island Press.

Srinivasan, V., and T.P. Clement. 2008. Analytical solutions for sequentially coupled one-dimensional reactive transport problems—Part I: Mathematical derivations. *Advances in Water Resources* 31, no. 2: 203–218.

Wiedemeier, T.H., H.S. Rifai, C.J. Newell, and J.T. Wilson. 1999. *Natural Attenuation of Fuels and Chlorinated Solvents in the Subsurface*, 617. New York: John Wiley & Sons.

Yu, S., M.E. Dolan, and L. Semprini. 2005. Kinetics and inhibition of reductive dechlorination of chlorinated ethylenes by two different mixed cultures. *Environmental Science & Technology* 39: 195–205.